**Deskripsi Tugas Besar MK Pembelajaran Mesin**

**2021**

Ketentuan Umum:

1. Tugas besar dikerjakan dalam dua tahap.
   1. Tahap pertama dilakukan secara **individu,**
   2. Sedangkan tahapkeduadilaksanakan secara **berkelompok maksimal 2 orang**.
2. Harus mematuhi etika plagiasi baik ide, kode, maupun laporan. Segala bentuk pelanggaran akan diberikan sanksi berat, minimal E untuk mata kuliah pembelajaran mesin
3. Deadline pengumpulan tugas tahap pertama adalah hari Jumat pada minggu ke-8 (pada tanggal 19 April 2021 jam 21.00 WIB), sedangkan tugas tahap kedua adalah pada minggu ke-11 (pada tanggal 8 Mei 2021 jam 21.00 WIB).
4. Presentasi tugas besar tahap pertama dilakukan secara individu kepada asisten, sedangkan presentasi tugas besar kedua dilakukan pada minggu 3 minggu terakhir perkuliahan di depan kelas dengan dosen pengampu.
5. Tugas dikumpulkan melalui LMS pada kelas masing-masing.

Deskripsi Tugas:

1. Setiap mahasiswa **harus** mengerjakan **2 buah task** terhadap dataset tersebut, yaitu **Clustering** dan **Classification**.
2. Untuk setiap task, mahasiswa **diharuskan** setidaknya melakukan hal-hal berikut:
   1. Formulasi Masalah: Jelaskan apa permasalahan yang ingin diselesaikan.
   2. Eksplorasi dan Persiapan Data (termasuk split data): Lakukan semua Teknik eksplorasi dan persiapan data yang menurut anda perlu dilakukan. Jelaskan mengapa perlu melakukan Teknik tersebut, dan lakukan analisis terhadap hasilnya.
   3. Pemodelan: Bangunlah model menggunakan data hasil proses 2.b, dan lakukan proses training untuk mendapatkan hasil terbaik. Jelaskan dengan detail semua proses yang anda lakukan, justifikasi, beserta analisis hasilnya.
   4. Evaluasi: Pilih metode evaluasi yang sesuai beserta justifikasinya. Lakukan evaluasi terhadap model yang telah dihasilkan. Berikan analisis terhadap hasil evaluasi.
   5. Eksperimen: Lakukan berbagai eksperimen yang melibatkan tahapan Eksplorasi dan Persiapan Data, Pemodelan, dan Evaluasi dalam rangka mendapatkan hasil terbaik. Laporkan semua Eksperimen yang dilakukan beserta analisis hasil dan perbandingannya.
   6. Kesimpulan: Sampaikan kesimpulan dari semua proses yang dijalankan beserta hasil akhir dari berbagai eksperimen yang telah dilakukan.
3. Untuk Task Clustering, mahasiswa **tidak boleh menggunakan library** untuk proses pemodelan (harus coding from scratch).
4. Bahasa pemrograman yang digunakan adalah **Python**.
5. Delivery tugas berupa:
   1. Laporan diketik dalam A4 menggunakan times new roman size 12, margin 3cm, space 1.15 (.pdf)
   2. Kode-kode yang siap dieksekusi (.py/.ipynb)
   3. Video presentasi dengan durasi 7-10 menit yang diunggah pada akun youtube. Sertakan tautan (link) pada laporan pada poin (a)
   4. Semua Dataset hasil setiap tahapan eksplorasi harus dicantumkan pada laporan poin (a)

Penilaian:

1. Laporan (30%):
   1. Kelengkapan bagian laporan (5%)
   2. Akurasi laporan terhadap yang dikerjakan/kode (5%)
   3. Analisis pemilihan Teknik yang digunakan beserta interpretasi hasil (20%)
2. Kode (40%):
   1. Ketepatan solusi beserta penjelasanya. Kode wajib diberikan komentar sebaik-baiknya dengan tujuan mudah dipahami oleh orang lain (evaluator)
3. Presentasi (30%):
   1. Organisasi konten presentasi (5%)
   2. Pemahaman tentang project (25%)

Dataset:

Dataset dapat diunduh menggunakan akun **Telkom University** di: http://bit.ly/dataML2021